International Journal of Science Education

ISSN: 0950-0693 (Print) 1464-5289 (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Teachers' use of learning progression-based formative assessment to inform teachers' instructional adjustment: a case study of two physics teachers' instruction

Xiaoming Zhai, Min Li & Yuying Guo

To cite this article: Xiaoming Zhai, Min Li & Yuying Guo (2018) Teachers' use of learning progression-based formative assessment to inform teachers' instructional adjustment: a case study of two physics teachers' instruction, International Journal of Science Education, 40:15, 1832-1856, DOI: 10.1080/09500693.2018.1512772

To link to this article: https://doi.org/10.1080/09500693.2018.1512772

Published online: 14 Oct 2018.

Submit your article to this journal

Article views: 686

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=tsed20

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
2018, VOL. 40, NO. 15, 1832–1856
https://doi.org/10.1080/09500693.2018.1512772

Routledge
Taylor & Francis Group

Check for updates

Teachers' use of learning progression-based formative assessment to inform teachers' instructional adjustment: a case study of two physics teachers' instruction

Xiaoming Zhai [a], Min Li[b] and Yuying Guo[c]

[a]Graduate School of Education, Stanford University, Palo Alto, CA, USA; [b]College of Education, University of Washington, Seattle, WA, USA; [c]Department of Physics, Beijing Normal University, Beijing, People's Republic of China

ABSTRACT

To explore the affordances of learning progressions (LP) in support of teachers' formal and informal formative assessment (FFA and IFA) practices, we conducted a case study with a fine-grain-sized energy LP. The study first theorises an LP-based formative assessment model, which proposes the use of LPs in clarifying learning objectives, eliciting and interpreting students' understanding, and acting in instruction. Then, we examine multiple iterations of two high-school teachers' lesson plans and their enactment of the lessons to identify the instructional adjustments teachers made as part of the LP-based FFA and IFA after professional training. We found that both teachers refined the learning objectives and activities and most of their adjustments promoted students' learning progression; their adjustments have an interdependency between objectives and activities. Both made more adjustments in the IFA than in the FFA, but the alignment with the LP was higher in FFA than in IFA. Contrary to the researchers' expectations, both teachers perceived the LP as a content structure. However, in practice, both employed it as a reference to interpret students' responses by comparing these responses against their own expectations. The teachers also reported collecting assessment information during teacher-student interactive activities and using this material to infer the level of students' understanding in order to decide on the next instructional activities. Both teachers reported that the process of directly engaging with the LP formatively resulted in their having a much more nuanced sense of students' understanding when they revisited and altered the sequence of the learning activities.

ARTICLE HISTORY
Received 17 September 2017
Accepted 13 August 2018

KEYWORDS
Learning progression; formative assessment; instructional adjustment; physics teacher; lesson plan; energy; case study

Introduction

Learning progressions (LPs) have been conceptualised as 'descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic' (NRC, 2007; p. 219). LPs have understandably drawn great attention from science curriculum designers (NRC, 2012, 2013),

CONTACT Yuying Guo yyguo@bnu.edu.cn Department of Physics, Beijing Normal University, Beijing, People's Republic of China

© 2018 Informa UK Limited, trading as Taylor & Francis Group

assessment scholars and, especially, science teachers (Duncan & Hmelo-Silver, 2009), as they provide a means of locating the level of students' thinking, thus allowing teachers to track students' progression towards learning targets (Gunckel, Mohan, Covitt, & Anderson, 2012). In response to the information about students' progression provided by LPs, teachers can generate instructional moves to directly support students' incremental steps towards learning targets (Furtak, 2012). However, LPs, when functioning as a final product without any corresponding professional development (PD), rarely offer teachers concrete, actionable information about how to meet students' learning needs in order to support their progress along the LP (Furtak, 2012). Without such professional support, individual teachers must make sense of the LP and translate it into practice on their own.

Previous studies indicate that science teachers are capable of translating LPs into practice in the processes of formative assessment and that there are benefits to using LPs to interpret students' performances, inform classroom practices, and, subsequently promote students' learning. However, we found a very limited number of empirical studies explore how teachers utilise LP in the formative assessment and no studies that differentiate teachers' engagement with the LP in formal formative assessment (FFA) and informal formative assessment (IFA). We don't know how teachers' understanding of LP helps them interpret the information collected from FFA and IFA process. Nor do we know how their understanding supports transforming this information into practice through adjustments to their instruction. In addition, teachers who participated in prior studies were either offered intensive pedagogical support (Yin, Tomita, & Shavelson, 2014) or involved in co-construction of the LP with long-term PD (Furtak & Heredia, 2014), both of which require a great amount of investment regarding time and money. More important, the majority of science teachers do not have the opportunity to attend such enriched and long-term PD programmes, especially in some developing countries such as China, which has more than sixteen million teachers. Instead, common ways for them to access the practical knowledge acquired by using LPs is by attending short-term lectures, reading relevant materials, and practising by themselves. Even though these practices are popular among teachers and research has shown that short-term in-service programmes make sense to enhance teachers' practical knowledge (Clermont, Krajcik, & Borko, 1993), little is known about how teachers understand LPs and use them in practice after attending short-term training.

This case study is intended to address this knowledge gap by studying how two high school physics teachers (one novice and one experienced) use an LP in their formative assessments when they are provided limited support resources. We focused on one aspect of formative assessment, by involving in LP-based formative assessment to adjust instruction. Specifically, we aimed to unpack the type and nature of instructional adjustment that the teachers made, the ways they applied the LP when making these adjustments, and the underlying reasoning with respect to the LP used to justify their instructional decisions. We asked the following research questions:

(1) How do a novice teacher and an experienced teacher engage with an LP-based formative assessment to adjust their instruction?
(2) What role does the teachers' understanding of the LP play in their use of the LP-based formative assessment in instructional adjustment?

In order to address these research questions, we first articulate the concepts of LP-based formative assessment and then describe the steps involved when teachers make use of formative assessment information to adjust their instruction.

LP-Based formative assessment

Consistent with most definitions of formative assessment (Black & Wiliam, 1998; Shepard, 2009), we conceptualise such assessment as a *process* that requires the active engagement of both teachers and their students. Further, formative assessment is a *cyclical* process of clarifying and communicating learning objectives, eliciting information about students' performance, interpreting the information, and acting on the information (Ruiz-Primo & Li, 2013).

Two kinds of LP-based formative assessment

With respect to the formality and spontaneity of the process, Shavelson et al. (2008) distinguish between two types of formative assessment processes: formal formative assessment (FFA) and informal formative assessment (IFA). While the former describes the deliberate insertion of assessment events to detect and act on the gap between what students know and what they need to know, the latter refers to assessment activities that occur on-the-fly when teachable moments arise.

As LPs explicitly describe students' progression of thinking, many studies have investigated and documented their utility in formative assessment (e.g. Alonzo, 2011; Furtak, 2012). An LP can be a specific tool for assisting teachers in developing both FFA and IFA instruments, as well as for interpreting students' performance. For example, an LP-based FFA may be implemented prior to a lesson, when a teacher uses an LP-based assessment instrument to collect information about students' prior knowledge and then uses the information to plan a lesson to address students' learning needs pinpointed by the assessment. LP-based IFAs are less formally implemented, such as when a teacher anticipates and interprets students' thinking at any time during a lesson, based on the teacher's own understanding of the LP. For example, an IFA might occur in a pair-share discussion when students interact with their peers, and in response, the teacher facilitates the conversation by asking several funnelling questions. Admittedly, the distinction between LP-based FFAs and IFAs sometimes becomes blurred due to the complex and dynamic nature of formative assessment and how it is blended with instruction (Minstrell, Anderson, & Li, 2016). For instance, a teacher may become aware of gaps in her students' thinking, based on an LP-based FFA prompt. With this understanding, she is alert to and anticipatory of such gaps while she participates in the whole-class discussion. She then engages with an LP-based IFA by asking a chain of probing questions that can precisely gauge where students are in the LP landscape and decides on a subsequent instructional move.

Unlike earlier studies that investigate how teachers and students benefit from rich assessment information gathered from LP-based formative assessments (e.g. Furtak, 2012), the present study focuses on *how* teachers act on LP-based assessment information when adjusting instruction, both formally and informally. More specifically, this case study is intended to provide detailed accounts of how teachers, guided by an LP framework, incorporate information about their students' performance, both when revising their lesson plans

and enacting such plans. We detail how teachers make these adjustments to both learning objectives and learning activities as a part of formative assessment practice, and how their understanding of LPs plays a role in these formative assessment practices.

Stages of LP-based formative assessment

We depict the definition of LP-based formative assessment by substantiating the four-stage formative assessment model proposed by Ruiz-Primo and her colleagues (e.g. Ruiz-Primo & Furtak, 2007; Ruiz-Primo & Li, 2013) in the LP context, with the LP in the centre, mediating all four stages (see Figure 1). Mager (1984) outlines three instructional plan questions—*where to go*, *how to get there*, and *where we are now*—that describe how instructional adjustment takes place. Mager's three questions correspond to the four stages of our model (i.e. stage 1, stages 2 and 3, and stage 4 respectively) and help clarify the relationship of Stages 1 and 4 with Stages 2 and 3. Specifically, Stage 1 (i.e. clarifying learning objectives) maps to Mager's instructional question *where to go*, and Stage 4 (i.e. acting) maps to *how to get there*. Mager's questions clarify that the relation between these two stages and Stages 2 and 3, which map to *where we are now*, can be nonlinear and interdependent instead of being a strictly linearly sequential process. Specifically, by offering detailed information about common student ideas, LPs can inform teachers' development and selection of assessment tools and strategies to elicit student ideas (Stage 2) and their interpretations of the elicited ideas (Stage 3). Both elicitation and interpretation of student ideas must serve the purpose of informing teachers' practices at Stages 1 and 4 and ultimately, improving student learning.

Our model highlights LPs as cognitive tools that teachers can apply at all four stages of the formative assessment process, in particular for making instructional adjustments. Learning objectives, which are the expectations that teachers set for students' learning and which direct learners *where to go*, often are and should be guided by national, state, and district standards. These learning objectives are *clarified and communicated* at Stage 1 of the formative assessment process. Based on what teachers learn via information from LP-based assessment data about their students' current thinking, teachers can expand, refine, or collapse the planned learning objectives or include new learning objectives that are suitable for their students.

We assert that LPs, by offering rich information from the perspective of learners, can influence how teachers design and select learning activities for their students. This

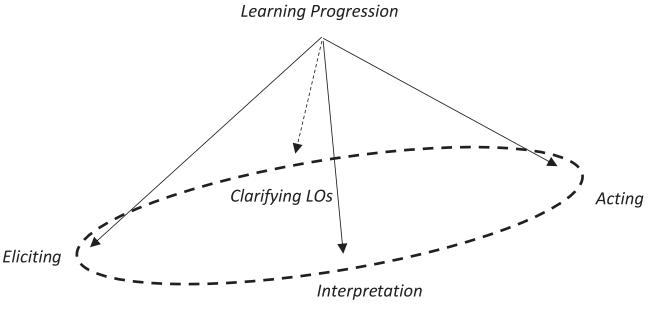

Figure 1. Model of learning progression-based formative assessment.

influence explicitly results in teachers' adjustment, either for planning lessons or for enacting lessons (Stage 4). This stage of instructional adjustment focuses not only on the learning activities—the building blocks that provide scaffolding to aid students' progression towards the learning objectives and the order or sequence of these blocks—but also the learning objectives that guide the formative assessment practice. Drawing upon formative assessment practices at Stages 1–3, teachers are able to diagnose students' thinking in an ongoing manner, identify the range and variety of ideas students hold, and arrive at and make informed decisions about what learning experiences need to be created. This is consistent with Margaret Heritage's observation that

> Only by keeping a very close eye on emerging learning through formative assessment can teachers be prospective, determining what is within the students' reach, and providing them experiences to support and extend learning through which students can then incorporate new learning into their developing schema. (2010, p. 8)

A productive formative assessment practice at Stage 4 requires teachers to translate LP-based assessment results from empirical information into actionable knowledge (Alonzo & Elby, 2015), based on both their understanding of the LP and their pedagogical expertise (Alonzo & Aschbacher, 2005; Gunckel, 2013). A teacher with a better understanding of the LP is expected to make better sense of the assessment information from both FFA and IFA, and more pedagogical knowledge means she/he is potentially able to act on this information more readily. This qualitative study aims to unpack the formative assessment processes of two physics teachers in order to understand how teachers utilise the LP perspective to convert assessment information into actions, as reflected in their adjustment of instructional purposes and moves.

Methods

Context

This study investigates two teachers' use of LP-based formative assessment from a large professional development (PD) project, in which 31 physics teachers from different schools participated in 64-hour training. These teachers attended lectures, observation sessions, discussion sessions, and workshops, and practised implementing LP-based formative assessments in physics lessons. The organiser, who funded this project, added a supplemental PD activity by asking the study trainees to mentor their colleagues through short-term workshops as well as discussion and introduction of PD materials, with the intention that the initial PD benefits could be expanded to the participants' colleagues. However, upon study completion, little was known about whether this workshop provided enough context regarding implementing LPs for it to be useful in their teaching practice, as LP-based PD generally requires a good understanding of LP, which requires a time-intensive process (Furtak, Morrison, & Kroog, 2014). To understand how teachers apply the LPs, if at all, in their formative assessments through this type of short-term PD workshop, we conducted the present exploratory study. We employed a theory-building approach (Eisenhardt, 1989), in which all the findings were supported by the four-stage LP-based formative assessment model. As in previous studies, we grounded this approach in a theory-related analytical generalisation rather than a statistical analytical generalisation (Yin, 2013).

Participants

Following Eisenhardt (1989), who stated that case studies should employ 'theoretical, not random, sampling' (p. 533), our sampling decision was driven by several theoretical concerns. Gunckel (2013) has recognised that teachers' 'pedagogical expertise' is necessary for them to apply LP in instruction effectively. Therefore, to offer a better understanding of how teachers perform in the LP-based formative assessment, it is necessary to take into consideration teaching experience, as it is regarded as the major source of teachers' pedagogical expertise (van Driel, Verloop, & de Vos, 1998). In addition, van Driel and his colleagues (1998) found that subject-matter knowledge is a predictor of science teachers' pedagogical performance. Thus, in the present study, we scaled by the teachers' preservice schooling background when we selected samples. This approach of sampling is not intended to construct a comparison between two contrary samples (i.e. novice and experienced), but to extend our understanding of the range of ways teachers implement LP-based formative assessment practice and why. Drawing upon the above concerns, we recruited two physics teachers (one novice and one experienced), and 78 Grade 10 students (approximately 15 years old) assigned to one of three classes taught by the two teachers in a high school in Beijing, China. At the time of the study, Nancy, a beginning teacher, had two years of physics teaching experience, while Eva, an experienced teacher, had more than 15 years of physics teaching experience. These two female teachers are certified in high school physics education, and Nancy holds a master's degree while Eva holds a bachelor's degree, both in physics. Nancy's background is typical of novice high school physics teachers in China, while Eva's background is typical of experienced teachers. The teachers had not used LP-based instruction or received any PD training about LPs prior to this study. Both teachers implemented their lessons using the same textbook (Zhang, 2006).

Procedures

This is a deliberately designed case study. The researchers first designed the FFA instrument as pre-test according to an established LP on energy and tested the students. The test result was used to trim the LP and to prepare training materials for the PD workshop. Before the workshop, the researchers pre-contacted the teachers and required them to prepare an initial version of lesson plans. The workshop took place three days before the formal teaching. Following the workshop, researchers held a three-day self-reflective session, during which teachers were required to reflect on their lesson plans and revise them. Teachers then used the revised lesson plans to enact the teaching in the classrooms. Researchers videotaped the teachings and administrated a posttest afterwards. Interviews were conducted with both teachers one week after the teaching. In the following section, we introduced the major study processes in detail.

Adapting the energy LP

In order to support the teachers' instructional adjustments based on the LP-based formative assessment, we adapted Wei's (2014) two-dimensional *Energy* LP. This LP was developed based on prior studies (e.g. Lee & Liu, 2010; Liu & McKeough, 2005; Neumann, Viering, Boone, & Fischer, 2013). Wei adapted five major constructs—*source*, *form*,

transformation, *degradation*, and *conservation* as the horizontal dimensions of the LP based on prior studies (e.g. Liu & McKeough, 2005; Neumann et al., 2013). However, unlike most prior studies that took integration or complexity of knowledge as a vertical variable, Wei derived reasoning approaches and principles as variables from two crosscutting scientific concepts: cause and effect: mechanism and explanation; system and system model. He claimed that the reasoning approaches denote the general ways of students' thinking (e.g. force-dynamic reasoning, mechanical reasoning and system-based reasoning; Talmy, 1985, 2000); the principles are those ideas that students use when explaining phenomena regarding energy (e.g. energy is human-centered; energy can be transported or flowing (Watts, 1983)). Both variables were adopted based on two facts: (1) the concept of energy was functionally created to explain observed phenomena that involve a system, even though energy is invisible to people; (2) students' increasingly sophisticated ways of thinking about energy are reflected by how they reason from the observed phenomena (effect) to the underlying mechanism (cause), and the principles they use during this reasoning process. Figure 2 presents an example that was extracted from Wei's LP (Wei, 2014, pp. 81–82). Specifically, the students in level 1 possess a force-dynamic reasoning approach, who perceive energy *as relevant to living organism, is used to fulfil the requirement of life, and is produced and consumed by living organism* (Watts, 1983); the students in level 3 with a mechanical reasoning approach possess the principle: *Energy is transforming when force does work, which indicates the unit of energy transformation*; while the students in level 5 employ a system-based reasoning, and they perceive energy *as a quantity characteristic of a system and a tool to help differentiate what might happen and what might not within a system*. After three rounds of verifications and adjustments with more than 3700 participants distributed from grade 7 to grade

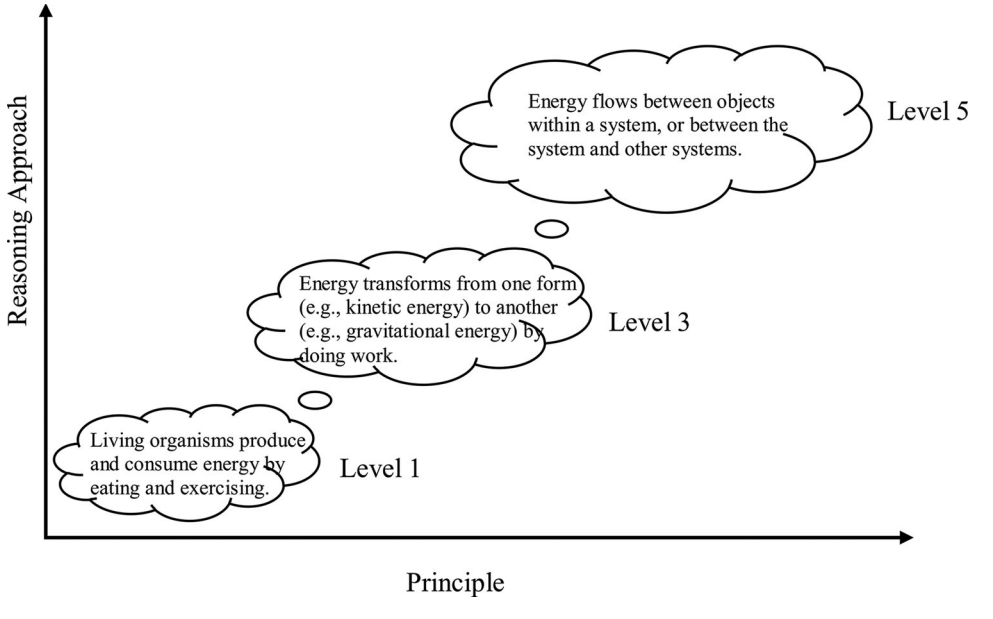
Reasoning Approach
Energy flows between objects within a system, or between the system and other systems.
Level 5
Energy transforms from one form (e.g., kinetic energy) to another (e.g., gravitational energy) by doing work.
Level 3
Living organisms produce and consume energy by eating and exercising.
Level 1
Principle

Figure 2. Wei's (2014) learning progression assumption: an example to illustrate the two progression variables.

12, he eventually developed five levels in the vertical dimension of the LP (Wei, 2014, pp. 111–112). In Wei's LP, the levels in different constructs are staggered, as displayed in other studies (Neumann et al., 2013; Wilson, 2009; Yao, Guo, & Neumann, 2017). For example, the levels in the *transformation* are more complex than those in the *form* so that the same level in *transformation* seems more challenging than that in the *form*.

This study focuses on the topic of *Work*, one section of the unit *Mechanical Energy and Energy Sources* for students in Grade 10. Work is a fundamental concept in the understanding of energy (Warren, 1982). Energy is not visualised, and it can only be measured when it is being transformed because work is being done (Lee & Liu, 2010). Thus, energy is commonly defined as the capacity to do work. Liu and McKeough (2005) found the understanding of work is a piece of foundational knowledge that provides access to higher level concepts such as energy transformation. In a review of textbooks and student interviews, we found students should have learned about simple forms of mechanical energy as well as the connotation of *work* within simple scenarios in Grade 8. However, the Grade 8 material mainly introduces how to calculate work instead of providing students with information about the specific connection between *work* and energy transformation (i.e. *work* is the magnitude of energy transformation). Nordine, Krajcik, and Fortus (2011) argued that

> many middle school energy curricula focus students' attention on performing simple calculations of energy quantities (e.g. work, kinetic energy, gravitational potential energy), such an emphasis necessarily limits the range of phenomena that students are equipped to understand and risks burying the central ideas of energy [e.g. energy transformations; Author added according to the context of the literature] in the details of calculation. (p. 695)

Therefore, the revised LP **in this study** was intended to expand students' understanding of *work* rather than the calculation of work, as it is connected to and extended the understanding of different forms of energy and to energy transformation.

We made three adaptions to Wei's original LP, which was validated by research that applied to students from Grades 7 through 12, such that the LP could be used in the current 10th-grade lesson on *work*. First, we selected from Wei's LP only the two major constructs in the horizontal dimension that are relevant to the topic of *work*: *form* and *transformation*. Second, after comparing against the requirement of the *Physics Curriculum Standard* in China, we realised that the 4th and 5th levels of the two selected major constructs in Wei's LP exceed the learning goals of the current lesson. Therefore, we removed these two levels and kept only the first three levels for each. Last, we refined the descriptions of those three levels by using practitioner-friendly language. For example, the initial 1st level of *form* was articulated as 'Students know about energy as relevant to some specific activity, but cannot identify the difference between force and energy'. We revised it as 'Students know about different kinds of energy and can identify whether objects possess energy for a given situation, but cannot describe the difference between energy and force'. The understanding of *work*, consistent with Wei's assumption, is staggered across all levels of the two constructs. To better understand the concept of *work*, students must first possess a robust understanding of *forms of energy* in a given level, after which they will be able to move on to the corresponding level of *transformation of energy*. Thus, students' understanding of *work* progresses along the levels of the LP between the two constructs (Zhai, Guo, & Zhang, 2015) (Table 1).

Professional development workshop

We selected 12 ordered multiple-choice items (Briggs, Alonzo, Schwab, & Wilson, 2006), within which there are three attached second-tier open-ended items, and one open-ended item from Wei's item pool to construct an FFA instrument. We then administrated the instrument in each teacher's classroom one week prior to the workshop. The Rasch analysis results and the qualitative summary of students' responses to the open-ended items for the entire sample comprised most of the PD materials, as they provided the teachers with relevant assessment information about the 78 students. A one-way between-subjects analysis of variance was conducted to compare the measures of student ability to understand energy from the FFA instrument for the three classes. There was no significant difference at the $p < .05$ level for the three classes [$F(2, 75) = 1.006$, $p = 0.370$]. We also developed a posttest, consisting of sixteen ordered multiple-choice items and one open-ended item, that was administered after the teaching. Both instruments were equated by eleven anchor items and students' ability measures, as estimated by a Rasch model, were compared by paired t-test for each class. The results suggest that after learning, each class had made significant progress at the $p < 0.01$ level (see Table 2). We note, however, that these findings are not the focus of this study and will be further explored in future studies; they are intended to function only as another predictor to judge whether teachers' engagement with the formative assessment is aligned with the *Energy* LP.

We organised a two-hour workshop session plus three one-day self-reflective sessions. The first author and another physics education researcher conducted the workshop with the two teachers. First, researchers presented and explained the description of the *Energy* LP to the teachers and answered their general and specific questions. Then, collaboratively with the two teachers, the researchers explored the results of the students' performance from the FFA instrument. These results were presented as (1) a Wright map (see Figure 3), which displays the distribution of both item difficulty and student ability on the same scale, to provide an overview of their students' performance; and (2) descriptions of each level of the LP along with the mapping of each option of the ordered multiple-choice items to these levels. We also provided teachers with information about the proportion of student responding at different levels of a particular item, as well as qualitative

Table 1. The adapted energy learning progression from Wei (2014).

Level	Forms of energy	Level	Transformation of energy
F3	Students are able to understand and analyse the ME[a] of a specific system, as well as the factors that influence the specific forms of ME quantitatively.	T3	Students understand the diverse phenomenon of energy transformation in a system and can use the concept of *work* to explain the transformation of ME in daily life.
F2	Students are familiar with different kinds of ME (e.g. KE[b] and GPE[c]), as well as the factors that influence their change, such as height for GPE.	T2	Students understand the diverse phenomenon of energy transformation in daily life and can explain the transformation of energy by 'force'; cannot connect *work* with the change or transformation of ME.
F1	Students know different kinds of ME; can identify whether objects possess energy in a given situation; cannot identify the difference between energy and force.	T1	Students know the phenomenon of energy transformation and can use an observable reason, such as eating food, crashing or flowing, to explain the transformation of ME.

[a]ME = Mechanical energy.
[b]KE = Kinetic energy.
[c]GPE = Gravitational potential energy.

Table 2. Paired t-test of student ability measures.

Class	Sample	Pre-test		Post-test		Gains	
		Mean	SD	Mean	SD	Mean	SD
Nancy	24	−0.46	0.63	0.92	0.90	1.39*	0.86
Eva	28	−0.41	0.64	1.10	0.75	1.52*	0.83
Eva	26	−0.23	0.57	1.56	0.67	1.79*	0.90

Note. *$p < .001$.

descriptions of typical student responses to the open-ended items. Teachers were asked to interpret the results according to their understanding of their own students. We expected that these resources would provide teachers with a comprehensive picture of their students' ideas about energy, and that, as described above, this information would inform teachers' revisions to their learning objectives and learning activities. During the workshop, teachers actively engaged in the discussion and offered their own understanding of the results. However, we did not provide specific strategies for utilising the LP or the results of the LP-based FFA.

After the workshop, teachers were asked to spend three days reviewing the materials and revising their lesson plans before the *Work* lesson; they were also given three additional resources: (1) instructions about how to use the workshop materials; (2) a practitioner article in Chinese that portrayed the instructional potentials of LPs, which

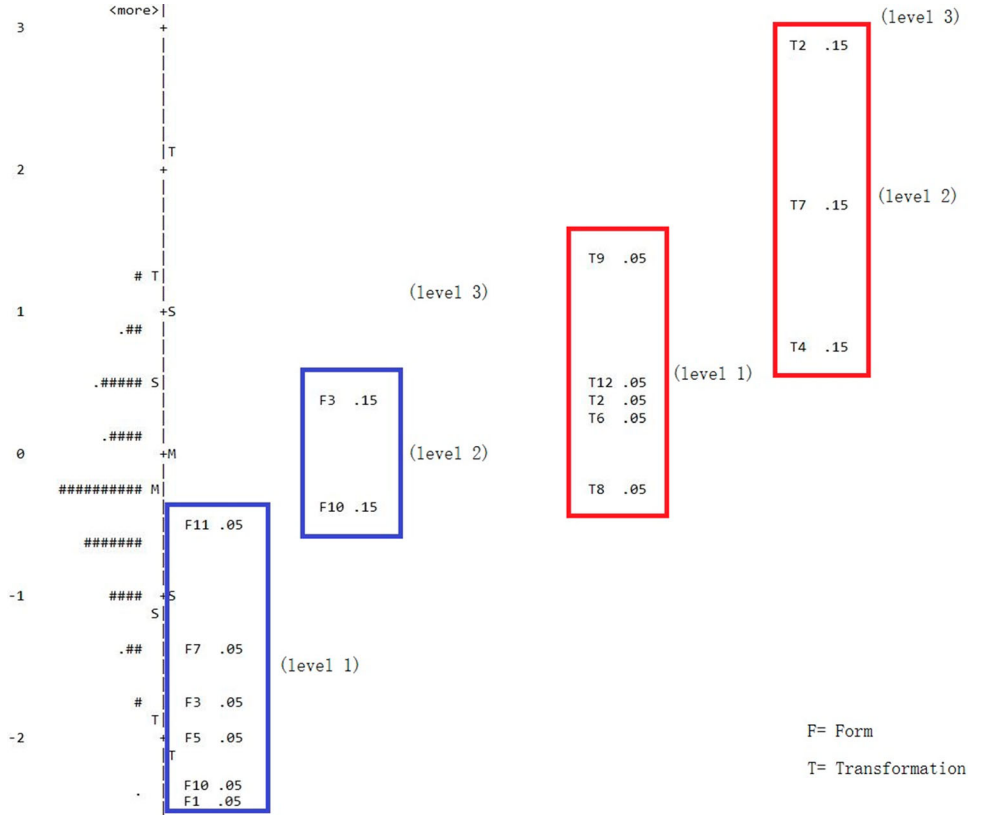

Figure 3. Wright map: students' understanding of energy form and energy transformation.

included classroom examples in the context of physics; and (3) examples from Wei (2014) for using the *Energy* LP to conduct instructional adjustment. In addition, we asked both teachers to review their lesson plans and to update them accordingly. Prompt questions such as 'In which way do you think this learning activity can facilitate students' progression? From where to where?' were also provided to teachers.

Data collection and analysis

Instructional adjustment

We collected teachers' lesson plans before the workshop (version I), asked the teachers to revise their lesson plans (resulting in version II, collected after the 3 one-day reflective sessions), and finally, videotaped their teaching of the lesson (the transcript lesson as version III) on the fourth day. We identified two rounds of revisions: (1) changes between version I and version II (in terms of learning objectives and learning activities) to determine how teachers revised their lesson plans that happened prior to the instruction informed by the FFA; and (2) differences between version II and version III (in terms of learning activities) to identify how teachers adjusted their planned lessons on-the-fly through the IFA.

By comparing the two versions of lesson plans and instructional enactment, we produced annotative notes to describe changes to the learning objectives or learning activities qualitatively (e.g. teacher's questions, student drawing activity) and assigned codes to each of these changes. For the learning objectives, coding included (1) content: *forms of energy* or *transformation of energy* and (2) nature of revision: minor editorial changes to the *wording* or more *substantive* changes. For the learning activities, our coding focused on three aspects: (1) content: *forms of energy* or *transformation of energy*; (2) type of revision: *adding* a new learning activity, *removing* a learning activity, or *adjusting the order* of the learning activities; and (3) form of the revised learning activities: *teacher-led* (e.g. teacher's explanation, demonstration), *student-led* (e.g. student's explanation, collection of evidence), or *teacher-student interactive* (e.g. teacher's question invites students to answer; teacher asks a follow-up question based on students' initial responses).

More important, for each of these changes, we also evaluated the quality of the revision in terms of how well the revision improves the instance's alignment with the LP and/or promotes students' progress: (1) *low* if a revision decreases the alignment of the instance (e.g. cognitive demand of a task is too high, given students' current LP level); (2) *medium* if a revision matches the alignment of the instance but does not increase the alignment; and (3) *high* if the revision increases the alignment of the instance (examples are presented in Table 3).

Together, the first author and a physics teacher consultant performed a consensus coding to identify and annotate the changes that had been made during the two rounds of revision. The researcher then assigned codes to these changes, and the teacher consultant (second coder) reviewed the codes and agreed in all but 17.7% of the cases. The two coders resolved all coding discrepancies through discussion.

Interviews

The first author conducted an audio-recorded semi-structured interview with each teacher one week after their teaching of the lesson. Interview questions were developed to capture each teacher's understanding of the LP, how the LP helped each teacher adjust the learning

Table 3. Examples of teachers' revisions for both learning objectives and learning activities.

Rev.	Ins.	Prior version	Post version	Coding note	Qua.
LOs	1	Cultivate students' learning disposition in the inquiry process. (Eva, version I)	***Connect the learning with phenomena in daily life in order to understand that transformation of energy is a general phenomenon around us.*** (Eva, version II)	The original LO was vague, while the updated one guides students to Level T2.	High
	2	Know the different energy forms, and that transformation is possible between these different energy forms. (Nancy, version I)	Know that different energy forms ***correspond to different states of objects*** and that transformation between different energy forms is possible. (Nancy, version II)	The revision makes the description more accurate and specific, but still remains in Level F1.	Medium
	3	Explore the law of nature from the perspective of work. (Eva, version I)	Explore the law of nature from the perspective of work; ***understand the concept of a system and its property in terms of how energy flows in or out***. (Eva, version II)	The revision exceeds the current highest-level T3, so it is not appropriate for the current students.	Low
LAs	4	T: Display: drop a box at different heights to hit a plasticine. T: Question: What kinds of energy do these objects have? (Eva, version II)	[T: Display: hold a box at a certain height T: What kinds of energy does the box have? S: Gravity potential energy] ***T: How do you know?*** S: It is at a certain height … (Eve, version III)	The question the teacher added led students to consider the factors that influence Gravity potential energy, and therefore indicates a high-quality revision by leading from Level F1 to F2.	High
	5	Question: A horizontal force F is applied to a stable object that is lying on a rough horizontal surface, and pulls the object along the direction of the force for a distance of x. How much work does the force do? (Nancy, version I)	Question: A horizontal force ***F = 10N*** is applied to a stable object lying on a rough horizontal surface, and pulls the object along the direction of the force for a distance of ***x = 5 m***. How much work does the force do? (Nancy, version II)	Adding numbers to these variables does not actually make sense in terms of directing students to a higher level.	Medium
	6	[T: Can you give examples to illustrate different energy?] (Nancy, version II)	[T: Please work with peers to generate examples illustrating different kinds of energy. S1: Motor consuming electric energy, battery consuming biological energy …] ***T: Where are these energies transformed from in each of your examples?*** [S1: The motor transformed electric energy to kinetic energy; the batter transformed biological energy into electric energy] ***T: What elicits the transformation?*** [S2: Maybe by doing work … ***T: Who does work? How?*** S: No response] (Nancy, version III)	The added questions were intended to promote students' progress from Level F1 to T3, but these additions are too difficult for students, especially as the questions lack anchor activities. Students S1 and S2 are high-performance students, while most of the students feel at a loss, according to the video.	Low

objectives and related learning activities in the LP-based formative assessment processes, and how each perceived the usefulness of the LP in the LP-based formative assessment. The interviewer sought to elicit teachers' descriptions of their practice by asking *how* questions and asking teachers to provide examples to illustrate their responses. For example,

we asked, 'How do you think the *Energy* LP helps support instructional adjustment?' followed by 'Please provide examples to illustrate your viewpoint'.

Interviews were transcribed prior to analysis. We first formulated their initial claims about the role of the LP in the LP-based formative assessment by reviewing and discussing the interview data. Second, we developed a coding scheme that included three dimensions: teachers' understanding of the LP, usages of the LP, and perceived usefulness of the LP, and then segmented the quotations and assigned a note to describe each quotation accordingly. Finally, we searched for evidence from the interview data to confirm or disconfirm teachers' claims corresponding to research question 2.

Results

Corresponding to our research questions, we report our results in two sections. In the first section, we present insights we gained regarding how teachers make their adjustments (the *action* stage in our theoretical model) towards learning objectives and learning activities that are explicitly displayed in the lesson plans and the teaching practice, as a result of the LP-based formative assessment practice. We also differentiate teachers' adjustments in FFA and IFA and the interdependency between adjustments towards learning objective and towards learning activity. By examining the above adjustments, we gain a better understanding of teachers' engagement in LP-based formative assessment. In the second section, we pinpoint what roles of teachers' understanding of LP plays in the FFA and IFA process. Findings from the interview data were triangulated by the adjustments we examined in the first section.

Teachers' engagement in the LP-based formative assessment

Revisions to learning objectives

The learning objective is regarded as similar to a destination in 'learning journey', during which teachers need to be very clear about where to go (Alonzo, 2011). A proper learning objective in a lesson plan is determined by two critical components: the expectations from the curriculum and the students' pre-understanding. Therefore, teachers have to evaluate and adjust the learning objective according to these two components. This process is especially difficult for teachers who are less knowledgeable about students' pre-understanding and have less pedagogical knowledge (John, 2006). LPs that are encapsulated in the FFA provide detailed information about students' baseline knowledge before learning and are expected to assist teachers in better adjusting the learning objective.

Our empirical results indicate that both teachers revised a relatively large number of learning objectives (7 out of 12 for each instructor) after receiving results from the FFA instrument. The revisions mostly involved substantive changes (78.6%), rather than merely small editorial changes; Nancy made such changes in 6 of 7 revisions, while Eva did so in 5 of 7. For example, Nancy, the novice teacher, added as a new takeaway that 'Work is the magnitude of the transformation of energy'. She also revised one stumbling block of learning, changing it from 'Using the mathematical formula of work' to 'Understanding the relationship between work and transformation of energy'. These changes seemed substantial since the former highlights the mathematical representation of work, whereas the latter focuses on the understanding of its connotation and settles work into energy. The experienced teacher, Eva, also primarily made substantive

adjustments based on information from the FFA instrument, although two of her adjustments to the takeaways seemed less substantial. The purpose of these changes appeared to be primarily to make the descriptions more precise. Most of the changes were substantive suggests that the LP-based FFA result informed both teachers with valuable information that they did not know or notice before the workshop.

Another predictor to judge whether teachers forwardly engaged with the LP-based formative assessment is the alignment of their adjustment—in other words, the quality of the adjustment—with the Energy LP. This empirical result indicates that most of the revisions both teachers made were judged to be at the medium level or above (i.e. appropriate for the students' LP level); a higher percentage of Eva's revisions were found to be high-quality (85.7%), as compared with Nancy's (57.1%). The high-quality revisions made the learning objectives more concrete and led to higher-level learning. For example, Eva changed a vague learning objective that seems unmanageable for teachers in Instance 1 (see Table 3). The revised learning objective was concrete to the understanding of specific physics concepts and corresponded to Level T2, the purpose of which is to require students to understand the diverse phenomenon of energy transformation in daily life. Considering that students were mostly in Level F1 prior to the lesson, this suggests that the revision increased the quality of the learning objective (examples of medium- and low-quality revision, refer to Table 3). These results indicate that both teachers deeply engaged with the LP-based FFA practice—the adjustments from collecting information to establishing a proper learning objective.

Revisions to learning activities

In the LP-based formative assessment, the adjustments of learning activities are another kind of action (Stage 4, refer to the theoretical model) that teachers take to respond to the assessment information they obtain, which can be efficiently examined in teachers' lesson plans and practice. Though LP is regarded as a useful tool to help teachers scale the assessment information in formative assessment practice (Alonzo, 2011), how teachers act on FFA and IFA still depends on the interpretability of the assessment information, and the sources or utilities that teachers were able to employ in response (Black & Wiliam, 2009). One significant difference between FFA and IFA is the amount and characteristics of assessment information the teachers are able to collect. As IFA is a fast-paced, interactive process with students, more moment-to-moment and spontaneous information is provided and therefore more adjustments are needed to follow the pace of the class, compared to FFA (Ruiz-Primo & Furtak, 2007). Our finding indicates both teachers made more on-the-fly revisions through the IFA (second round, moving from version II to version III, $N = 102$) than revisions through the FFA (first round, moving from the version I to version II, $N = 26$). This indicates that no matter their teaching experience, the two teachers were more skilled in IFA than in FFA.

The second difference between FFA and IFA is that only in the former do teachers have the opportunity to acquire help from other sources (e.g. teaching books, consulting colleagues) when their pedagogical level of expertise does not meet their needs. Previous research (Heritage, 2007) has shown that due to teachers having the opportunity to seek help and more time to reflect on their adjustments within FFA, this type of assessment is maybe especially beneficial to novice teachers, as they are likely to need more external help compared to experienced teachers. In contrast, the moment-to-moment IFA does not allow teachers to engage in either activity (Heritage, 2007). Our empirical

results support these findings that level of expertise affects the way teachers implement assessments, as Nancy made more than twice as many revisions (18) as did Eva (8) in response to the FFA, but fewer revisions (39) than did Eva (63) in response to the IFA. Nancy's increased activity in the FFA suggests she may have had more learning activities that needed to be revised than Eva and these revisions were accessible, whereas for the IFA, she might have fewer activities to revise than Eva did, or she had inadequate opportunity and resources to make the necessary revisions.

A closer look at the quality of both teachers' adjustments provides a clearer understanding of their performance in the formative assessment. As stated earlier, we judged the quality of the revisions to learning activities based on whether the revised activities led to higher-level learning that consistent with the LP. A higher-quality revision indicates that teachers are not only capable of eliciting and interpreting the information, but also able to act on this information. We illustrate examples of revision quality (from low to high) in Table 3. Instance 4 that happened in the IFA is an example of a high-quality revision, in which Eva added an on-the-fly question to lead students to think about the factors that influence gravity potential energy, that led students from Level F1 to Level F2, the latter of which is defined as 'Students are familiar with different kinds of mechanical energy, as well as the factors that influence their change'.

Whereas revisions designated high-quality are those that promote learning from a lower level to a higher level, those designated medium- and low-quality revisions fail in this function. The designation of a revision as medium quality indicates that the adjustment did not promote learning from a lower level to a higher level, but also did not cause a decline in learning. For example, Instance 5 in Table 3 displayed Nancy's revisions in the FFA. She added values to these variables, but the changes did not lead to higher-level learning, nor did it precipitate a decline in learning. The low-quality revisions are those activities that were adjusted to a level too challenging for the students' current level and consequently led to a decline in their learning. As evidenced in Table 3, Instance 6, the last question the teacher proposed, 'Who does the work? How?', overstepped students' current learning level and led to a decline in learning progress. The initial question tapped into students' understanding of the energy form (Level F1), whereas the revised version, which failed in its purpose, intended to encourage students to 'use the concept of *work* to explain the transformation of mechanical energy' (Level T3). In such cases, the revisions are regarded as of low quality.

In sum, high-quality revisions comprised the largest percentage of revisions for both teachers (50.0% for Nancy and 62.5% for Eva) in FFA, with less than 5.6% judged to be of low quality. In IFA, compared to FFA, the percentage of high-quality revisions for Nancy decreases to 41% and the low-level revisions increased to 20.5%; Eva's revisions slightly decline in quality (61.9% are high-level). In all, the average quality of teachers' revisions has a slight drop in IFA compared to FFA.

Interdependency between adjustments towards the learning objective and the learning activity

The **adjustments** towards the learning objective and learning activity are regarded as ways to act on the assessment data, as part of the LP-based formative assessment practice (Stage 4). As pointed out earlier, clarifying the learning objective and acting (mostly towards learning activities) are two interdependent stages in the formative assessment

process, although they take place at two different stages. In this sense, teachers' adjustments to the learning objectives potentially impact their adjustments to their learning activities, functioning as a formative assessment circle, likely resulting in a resonance process.

Two observations provide insight into the resonance characteristics of the LP-based formative assessment. First, the correspondence between the adjustments to the activities and to the learning objectives was found for both teachers in FFA (not available in IFA due to the lack of explicit adjustments to learning objectives). For example, Eva added a new objective as 'Being able to analyse the energy change of a system in a physical process'. Two of her adjustments to learning activity target to this adjustment: In one place, she added an activity that requires students to generate examples from life that demonstrate energy change; in another place, she added a teacher demonstration including three processes (at rest, moving horizontally, and moving upwards) for a block that is subjected to a spring force, and required students to analyze the energy change. We found that 22 of the 26 adjustments to the learning activities in FFA for both teachers have such correspondence, with only four exceptions, all of which were Nancy's adjustments. This high ratio of correspondence indicates a high resonance between the two kinds of adjustments, which occurred even though the teachers were not explicitly instructed that the adjustments to the activities should be based on the adjustments to the objectives. The four out-of-correspondence instances are likely due to a lack of such focus on the learning objectives. For example, Nancy presented a picture in which an athlete is lifting a weight and then required students to 'Analyze the type(s) of energy the weight might possess'. However, this activity is not aligned to any of the learning objectives in her lesson plans.

Second, the correlation between the alignments of the LP with the adjustments to the learning objectives and to the learning activities helps us better understand how well the teachers' adjustments to learning objectives and to activities happened simultaneously. For the above example, the new objective Eva added is regarded as high quality because it maps to level T1, 'Students know the phenomenon of energy transformation and can use an observable reason, … to explain the transformation of mechanical energy', and most students were found to be below this level in our pre-test. For the newly added activities, the first one was regarded as not promoting students' learning progression (e.g. medium quality) because in the previous section teachers had demonstrated similar examples (e.g. throwing a ball; lifting a box). However, the second added activity is regarded as high-quality because it introduces a new form of mechanical energy—elastic potential energy that leads students to progress from level F1 to level F2—'Students are familiar with different kinds of mechanical energy … '. Ideally, we expect teachers' adjustments to the objectives will align with those made to the activities in terms of alignment with the LP (quality) if teachers perform well on the formative assessment. To assess this proposal, we conducted a Kendall correlation analysis (exclude missing value) between the two kinds of qualities and found a significant correlation (r_τ (22) = .43, $p < .05$). These results provide further evidence that high-quality adjustments may result when adjustments to the learning activities are interdependent with those made to the learning activities in the LP-based FFA practice.

The roles of the LP in formative assessment

Teachers' understanding of the LP

According to the interviews, both teachers appeared to consider the LP to be a content structure (Shavelson & Kurpius, 2012), in other words, based on a content rationale rather than having a focus on student thinking. Eva indicated that the LP was aligned to some degree with her own understanding of a content map of energy. After the workshop, she stated 'I did not feel any surprise about the LP'. She explained that she used to plan her own teaching by aligning it with an 'LP': 'For example, I used to first introduce the connotation of work, then introduce its formula, and then its negative and positive … all of these follow a rationale, and should be an LP, aha, am I right?' She also referred to 'a chapter perspective' many times to illustrate how she envisioned the similarity between the LPs and the structure of the content in the textbook and highlighted that the rationale under the content components was like an LP from her own perspective. Based on these interviews, we inferred that Eva had a different understanding of the LPs from what the researchers initially intended. In addition, perhaps because she is an experienced teacher, she thought the LP already reflected the established content structure in her mind and referred to this structure to explain why her initial instructional adjustment required less revision during the FFA.

Similarly, Nancy described the levels of the LP using an analogy of 'the counterpart knowledge students had learned in middle school' and 'the knowledge they are learning currently'. For example, the 8^{th}-grade textbook introduces how to calculate work by formula when the force is aligned with the motion, whereas the 10^{th}-grade textbook is organised at a higher level—demonstrates how to calculate work by formula when force is not aligned with motion. These refer to the content structure organised by the textbook, which was organised and sequenced by experts according to content rationale rather than students' actual understanding reflected in the empirically tested LP. Thus, we infer that Nancy also considered the LP to be a type of content structure. However, she indicated that she often has to refer back to the 'textbooks, reference books', or her mentor to help her set learning objectives or learning activities. Therefore, we suspect that she may not yet have established a content structure in her mind, which is reasonable for a novice teacher, especially compared to Eva, and this may have prompted Nancy to make more revisions on learning activities than Eva did (18 vs. 8) during the FFA where she had the opportunity to seek additional resources and support, whereas the pattern was reversed in the IFA (39 vs. 63, respectively) when she had to adjust her instruction on the fly without any external help.

Ways teachers utilised the LP

The teachers reported using the LP as a reference in three ways: (1) to interpret students' responses, (2) to infer students' understandings, and (3) to inform the sequence of their learning activities. First, the teachers described using the LP as a reference to interpret students' responses to the FFA instrument. Teachers used the interpretation by comparing it with their initial expectations of student understanding. Whereas Eva already had realistic expectations of her students' performance—she 'really did not expect they had a deep understanding prior to the instruction', and the Wright map confirmed [her] viewpoint, Nancy's expectations exceeded her students' understanding level, as reflected in their

responses to the FFA instrument. Nancy explains her use of the LP: 'I realized that my expectations of my students' performance were initially too high after I saw the results … Yeah, the LP makes clear the distance between my expectations to their actual understanding'. Both teachers made several references to the capacity of the Wright map to explicitly locate students' level of understanding. According to Eva, 'the LP [displayed in the Wright map] looks like a summary of the test results, which is more visual than students' scores for exploring students' proficiency of understandings'. From conventionally displayed assessment results, the teachers mainly got a total score or result describing only right/wrong, which lacked information that quantified the distance from students' current understanding to the learning objectives. The LP, however, by explicitly informing teachers about the extent to which their students understand the concepts, indicates the gap in terms of the number of levels between students' current understanding and the learning objectives. Therefore, it is more interpretable when the goal is comparing students' prior understanding with the teacher's expectations.

Our findings indicate that these comparisons directly informed teachers' revisions of the learning objectives. For example, consider Instance 1 in Table 3, in which Nancy adjusted a vague learning objective to make it more accurate. She reported that the LP first brought her back 'on track' to teaching *work* by progression from the *forms of energy* to the *transformation of energy*. After reviewing the distribution of the students [mostly in Level T1], she decided to adjust the initial vague learning objective into Level T2. 'The LP reminded me of the gap between my expectations and my students' actual proficiency level of understanding', Nancy stated.

Eva reported that another way of using the LP to interpret students' performance in the IFA is to collect assessment information from interactions with students during teacher-student interactive activities: 'At a glance, the LP provided information about the students' progress in understanding *work* … while the teacher-student interactive activities provided information about where students' understanding lies [in the LP], and ultimately helped me adjust the pace of my teaching'. She added many new teacher-student interactive learning activities (46% as compared to Nancy's 28%), which she believed would elicit students' understanding during her interaction with them. She provided an example to illustrate (Instance 4 in Table 3). Initially, Eva only planned to help students understand the *form of energy* by a series experiment demonstrations [Level F1]. After collecting information from a newly added question, 'how do you know [the gravity potential energy]?', she realised that students' understanding had already progressed to Level F2 and she no longer needed to run the demonstration of dropping a box from different heights, which was intended to show that an object at a certain height owns gravity potential energy; instead, she asked more challenging questions that led students to understand energy transformation in Level T2.

Second, teachers used the LP to infer students' understandings of the material. Eva used the levels of the LP to infer students' ideas from one level to another. For example, she planned a learning activity in which, in versions III, she held a box at a certain height and asked questions about the energy the box possesses (see Instance 4 in Table 3), but she did not ask students to analyse the energy transformation in the process of dropping the box. She explained how the LP informed an inference about student thinking that led to this decision: '[Transformation of energy is impossible for students to handle] such as in the process of dropping a box; some students do not even understand why gravitational potential energy decreases while kinetic energy increases at Level F2, so they definitely

do not know the quantitative relationship of the change of the energy that is described in Level T3'. Maybe due to her realisation that students were mainly on Level F1, she added many more activities on *forms of energy* compared to Nancy, especially in the IFA (16 compared to Nancy's 8).

Third, the LP helped teachers to sequence their learning activities. Both teachers made significant revisions to their instructional sequences by adding, deleting, and adjusting the order of activities during the FFA and IFA. Eva used an example to illustrate:

> Given the performance of students [on the FFA instrument], I planned my sequence from the phenomenon [experiment] of *forms of energy* [Level F1], which will naturally lead to knowledge about the influence factor for *energy* [Level T2]; the more challenging task is to link *work* to *transformation of energy* [Level T3]. Therefore, students must learn the concept of work as well as how to calculate work and what is positive or negative work before reaching Level T3. I planned my activities along this thread.

To pursue this alignment, Eva added the experiment in version II; in version III, she added teacher-student interactive questions to elicit students' understanding of the influential factors for mechanical energy and deleted the repetitive calculation activities. She also adjusted the order of activities in version III. For example, she initially planned to play a video showing toxophily and then ask students to analyse the energy transformation process, but decided to reverse the order. 'Since I knew that students already knew the influential factor of mechanical energy [Level F2], I asked them to think by themselves first, and then I played the video'. This comment indicates Eva tried to sequence her learning activities around the LP level that her students showed.

The perceived usefulness of LPs

The teachers' self-report indicated differences in perceived usefulness of LPs. Eva seemed to see the LP as having more direct potential to improve her teaching, as compared to Nancy. When we asked the teachers to evaluate the utility of the LP, Eva thought it could make instruction more effective:

> If teachers can relate the description of the LP to their instructional adjustment, it [the LP] really makes sense … The reason why I paid more attention to *forms of energy* than before was partial because of the LP, especially when I saw the students' distribution in the lower levels, and I appreciate the way the results of students' responses were displayed in the LP.

At the same time, she did not intend to rely solely on the LP: 'The LP is only one aspect with which I am concerned. The more important concern is my own experience'. Eva is a teacher with more than 15 years of teaching experience. Therefore, she was able to speculate about students' prior understanding when making instructional decisions according to experience: 'I guessed that students would experience difficulties here, as my former students did'. She provided an example:

> Although the theme is *work*, I don't think the key is *work* from a whole-chapter perspective. *Forms of energy* is the background of the entire chapter. If we only focus on *work*, it will lead students to pay more attention to the formula of *work* rather than the scientific connotation, which is inefficient for students … I would know about this even if I did not recognise it in the LP.

Eva's experience informed her to 'not only focus on *work*', but the LP reminded her to pay more attention to this. From Eva's self-report, we can see that the LP was useful due to its being (at least partially) consistent with, and eventually confirming, her prior conceptions.

In contrast, perhaps because Nancy lacked Eva's teaching experience, the LP provided more benefits to her than to Eva when interpreting students' performance and identifying students' learning gaps, leading her to a deeper consideration of her instructional adjustments. Nancy stated, 'The LP, in my mind, is a tool for a teacher to monitor the class. I feel it makes sense to use this tool, especially when I am puzzled by students' responses'. What made a deep impression on her was Instance 6 (Table 3), in which there was no response from students after her question. She reflected, 'If I could have connected students' responses with the LP earlier, I might not have asked that question'. Although Nancy realised the potential of the LP, to translate an LP-based diagnosis into practical knowledge in an on-the-fly situation requires more than just information about student understanding, and this may be challenging for novice teachers. We did find that the LP helped her refine her instructional objectives and activities in the FFA significantly, but the potential in the IFA for Nancy might be amplified if she were provided sufficient anchor strategies from the PD or allowed external support.

Discussion

Even though the potential of LPs was recognised for science education, there are a very limited number of studies (6 out of 26) investigate teachers' learning and teaching practice according to a most recent review study by Jin and her colleagues (2017). The current study adds to the literature by exploring how an experienced teacher and a novice teacher conducted instructional adjustments in LP-based formative assessment processes after a short workshop. The workshop deliberately provided information about students' prior understanding of *work* in an *Energy* LP and followed with a guided self-reflective session to support teachers in applying their understanding of the LP to planning and enacting lessons. We found many interesting findings:

As illustrated in our theoretical model, findings indicate that both teachers adjusted their learning objectives and learning activities as a result of the LP-based formative assessment practice. Both teachers made substantive adjustments to their learning objectives, including takeaways and stumbling blocks in the LP-based FFA. For both teachers, more than half of the revisions they made were regarded as facilitative of students' learning progression, though Eva made more such revisions than Nancy did. Results also indicate that both teachers revised a large number of learning activities in both the FFA and IFA process. Among the adjustments, close to a half were regarded to be high quality. These adjustments indicate that both teachers were deeply involved in the FFA and IFA practices.

More specifically, our finding highlights the differences between LP-based FFA and IFA, which informs us how to better accommodate the two kinds of formative assessments. First, LP-based IFA elicits more real-time information than FFA, and this information is a great source for teachers to elicit, interpret, infer, and make adjustments according to their understanding of the LP, as teachers obtained continuous and immediate feedback from their students. Thus, teachers made more adjustments in LP-based IFA compared to LP-based FFA. Second, because the LP-based FFA allows time for teachers to reflect before refining their instructional adjustments and allows teachers to seek external help when they meet pedagogical dilemmas, the LP-based FFA may be more accessible, especially to novice teachers. This implication was evidenced by the quality of teachers' adjustments—a higher alignment with the LP in FFA than in IFA, on average. This

suggests more support is needed in the LP-based IFA, compared to FFA, in order to further the implementation of LP-based formative assessment. A useful focus for future research would be to examine what specific assistance different teachers need in real-time LP-based IFA.

This study also provides evidence that the formative assessment stages have an interdependency relationship. However, our results demonstrate evidence for this relationship between the first (i.e. learning objectives) and fourth (i.e. action) stages only, as Stage 2 (i.e. eliciting) and Stage 3 (i.e. interpretation) are intrinsic and therefore difficult to capture directly. Thus, teachers appear to be concerned with, or at least impacted by, the learning objectives when adjusting the learning activities and vice versus, even though they may not be instructed to consider this interrelationship intentionally. This result suggests that in practice, the stages of the formative assessment are a holistic practice, and any effort intended to promote teachers' performance on LP-based formative assessments should consider these stages simultaneously.

We further identified three ways that teachers used the LP in their formative assessment practice. First, both teachers used the LP to interpret the results of the LP-based FFA instrument, locating students' understanding in comparison to their own expectations, and consequently refining the learning objectives. We assert that the LP may support such interpretations by summarising students' understanding and displaying it in different levels, especially within the Wright map. Presenting the FFA results as corresponding to LP levels was a more manageable format for teachers to interpret, compared to directly examining the students' responses or scores, and thus helped to transform information about student understanding into a form useful for instructional adjustment. This approach goes beyond the typical 'get it or don't' interpretation (Otero & Nathan, 2008) by treating students' understanding developmentally. In addition, Eva also reported using the LP in combination with teacher-student interactive activities to elicit and interpret students' understanding in the IFA. For example, she compared the elicited performance with the levels of the LP to locate where students were, and consequently adjusted her teaching.

The second way that the teachers used the LP was to infer students' understanding during IFA. Teachers often struggle with making instructional decisions following students' on-the-fly feedback, not only because they don't know the level of students' understanding, but they don't know what students will know after their next instructional act. LPs offer a means of providing this missing evidence for teachers so that they can infer from students' performance to students' probable feedback to a follow-up learning activity. Although Eva followed this progression, we did not find this use evidenced in Nancy's self-report. Nancy appeared to be overcautious and to lack sufficient pedagogical content knowledge critical to carrying out these steps (Haug & Ødegaard, 2015), which might be due to a lack of teaching experience. Because IFA requires teachers to elicit students' performances and to interpret and transform elicited information into practice immediately, the novice teacher, Nancy, might have felt more challenged. These findings suggest that level of experience may play a vital role in the ways teachers use LPs.

In Furtak and Heredia's study (2014), evidence suggested that teachers' involvement in co-developing and refining an LP may help them more successfully use the LP to sequence their instructional activities. Both teachers in this study used the LP directly to inform instructional sequences, even though they were not involved in the development of the LP. The LP, adapted to a classroom scale, may have aided in teachers' sense-making

and use, as it was similar in appearance to the content structure that teachers were familiar with. Thus, they may have been able to use the LP to sequence activities, even though they did not conceptualise the LP correctly. This indicates that an approach of organising short-term PD—intensive hands-on workshop plus self-reflective session makes sense to teachers regarding the implementation of a specific topic. To address the lack of long-term PD programmes on a specific topic, such as LP-based formative assessment in China, the study deliberately designed this short-term training, which is available for the majority of physics teachers. Unlike the prior medium- or long-term PD programmes in implementing LP that require significant time investment (Furtak & Heredia, 2014), the training in this study requires only a two-hour formal workshop. Though brief, this workshop resulted in both teachers finding ways to make sense of the LP in their formative assessment practices. This reminds us that the key issue might not be the length of the workshop itself, but how much the teachers take away from experience. The follow-up reflective session in this study is also of note, as it gave teachers more freedom to make sense of and practice with the material, and motivated them to reflect on student progress by providing support materials (e.g. the Wright map); this session was valuable because it showed promise as an alternative when traditional face-to-face training, which is more time- and cost-intensive, is not available. The accessibility of the training material is another major factor that contributes to the PD. In this study, the LP that was adapted to a classroom scale contributed to teachers' taking-out because it was more comprehensible and interpretable than prior large-grain LPs in common.

This study also documents the challenges teachers encountered when learning and implementing the LP-based FFA and IFA. First, both teachers did not understand the LP in the way the researchers expected. They both regarded the LP as a content structure rather than a reflection of students' cognitive development upon understanding the big idea of energy. This implies a need in teacher PD for addressing 'student-centred' pedagogy. Teachers need to internalise the LP into their pedagogy as a way of understanding their students rather than the content. Second, a lack of pedagogical content knowledge might be a barrier to better implementation of LP-based formative assessment, especially in on-the-fly formative assessment. For example, Nancy's adjustments remain a higher quality in FFA, compared to IFA, maybe because she could acquire external support in FFA, whereas in IFA, this advantage is absent. This challenge seems like more urgent for novice teachers than for experienced teachers. Consistent with Yin and her colleagues' (2014) suggestion, we suggest that further organised pedagogical support such as an LP-based curricular sequence might be helpful for teachers who lack the knowledge to infer students' performance.

Finally, we note some limitations to our work and outline some areas for future research. First, our theoretical model indicates that the LP mediates all four stages of formative assessment practice, but due to the research design, we only were able to examine the interdependent relationship between learning objectives and learning activities. Future studies should employ other methods to examine and verify the assumption of interdependency between all four stages. Second, this study finds a lack of pedagogical knowledge is potentially a barrier that challenges teachers' LP-based formative assessment practice, especially for novice teachers. It will be invaluable to determine what pedagogical knowledge is especially needed to support teachers' LP-based formative assessment practices. Future studies of required pedagogical knowledge should focus on both novice and

experienced teachers. Finally, even though we found that LP-based formative assessment is effective for teachers' instructional adjustments, we don't know whether this effect is fundamental or content-specific. Is the benefit from this specific workshop transferable to other topics? Investigating what features would make these knowledge transferable to other content would be a worthy question for future study.

Conclusions

This study specifies the usage of LP in both the FFA and IFA practice. The findings provide some promising indicators of how teachers may use an LP to interpret students' responses, infer students' understandings, and made instructional adjustments accordingly. The study also indicates an interdependency between the adjustments to learning objectives and the adjustments to learning activities. This suggests that clarify learning objectives, eliciting and integrating students' ideas and making instructional adjustments based on LP is a holistic process. Because IFA provides a large amount of rich information instantaneously, it can be challenging to integrate that information stream into practice, so teachers need more help in the IFA process than in the FFA process. The FFA is more accessible to teachers than IFA because during the process, teachers have more time to process the information and can look to external resources for help. To better implement LP in practice, teachers need to embed the LP into their pedagogical practices and internalise it as a way of understanding their students. A lack of pedagogical content knowledge is a barrier to implementing LP-based formative assessment. In addition, our results provide some preliminary evidence that incorporating reflective sessions with prompts to a short PD workshop may be useful and efficient in supporting teachers' use of LPs for formative assessment purposes.

Acknowledgement

The authors are grateful for the invaluable and insightful comments from Dr. Alicia C. Alonzo. Thanks to doctoral student Klint Kanopka who proofread the manuscript. The authors are grateful to the editors and reviewers for their comments and suggestions, which have enabled substantial improvements to be made to the article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by Advanced Innovation Center for Future Education, Beijing Normal University [grant number BJAICFE2016SR-005]; Ministry of Education of China [grant Number 13YJA880022]; Chinese Scholarship Council [grant number 201506040139].

ORCID

Xiaoming Zhai http://orcid.org/0000-0003-4519-1931

References

Alonzo, A. C. (2011). Learning progressions that support formative assessment practices. *Measurement: Interdisciplinary Research and Perspectives*, *9*(2–3), 124–129.

Alonzo, A. C., & Aschbacher, P. A. (2005, April). *Factors which contribute to teachers' adoption of science note- books as a formative assessment tool*. Paper presented at the annual meeting of the American Educational Research Association, Montréal, Quebec, Canada.

Alonzo, A. C., & Elby, A. (2015). *One physics teacher's use of a learning progression to generate knowledge about his students' understanding of force and motion*. Paper presented at the meeting of the American Educational Research Association, Chicago.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7–74.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (Formerly: Journal of Personnel Evaluation in Education)*, *21*(1), 5–31.

Briggs, D. C., Alonzo, A. C., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, *11*, 33–63.

Clermont, C. P., Krajcik, J. S., & Borko, H. (1993). The influence of an intensive in-service workshop on pedagogical content knowledge growth among novice chemical demonstrators. *Journal of Research in Science Teaching*, *30*, 21–43.

van Driel, J. H., Verloop, N., & de Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, *35*(6), 673–695.

Duncan, R. G., & Hmelo-Silver, C. E. (2009). Learning progressions: Aligning curriculum, instruction, and assessment. *Journal of Research in Science Teaching*, *46*(6), 606–609.

Eisenhardt, K. M. (1989). Building theories from case study research. *Academy of Management Review*, *14*(4), 532–550.

Furtak, E. M. (2012). Linking a learning progression for natural selection to teachers' enactment of formative assessment. *Journal of Research in Science Teaching*, *49*(9), 1181–1210.

Furtak, E. M., & Heredia, S. C. (2014). Exploring the influence of learning progressions in two teacher communities. *Journal of Research in Science Teaching*, *51*(8), 982–1020.

Furtak, E. M., Morrison, D., & Kroog, H. (2014). Investigating the link between learning progressions and classroom assessment. *Science Education*, *98*(4), 640–673.

Gunckel, K. (2013). *Teacher knowledge for using learning progressions in classroom instruction and assessment*. Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA.

Gunckel, K., Mohan, L., Covitt, B., & Anderson, C. (2012). Addressing challenges in developing learning progressions for environmental science literacy. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 39–75). Rotterdam, The Netherlands: Sense Publishers.

Haug, B. S., & Ødegaard, M. (2015). Formative assessment and teachers' sensitivity to student responses. *International Journal of Science Education*, *37*, 629–654.

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, *89*(2), 140–145.

Heritage, M. (2010). Formative assessment and next-generation assessment systems: Are we losing an opportunity? *Council of Chief State School Officers*, 1–20.

Jin, H., Mikeska, J., Hokayem, H., & Mavronikolas, E. (2017). *A Systematic Review of Learning Progression Research*. Paper presented at the National Association of Research in Science Teaching, San Antonio, US.

John, P. D. (2006). Lesson planning and the student teacher: Re-thinking the dominant model. *Journal of Curriculum Studies*, *38*(4), 483–498.

Lee, H., & Liu, O. L. (2010). Assessing learning progression of energy concepts across middle school grades: The knowledge integration perspective. *Science Education*, *94*, 665–688. doi:10.1002/sce.20382

Liu, X., & McKeough, A. (2005). Developmental growth in students' concept of energy: Analysis of selected items from the TIMSS database. *Journal of Research in Science Teaching*, *42*(5), 493–517.
Mager, R. F. (1984). *Preparing instructional objectives* (2nd ed.). Belmont, CA: Fearon-Pittman.
Minstrell, J., Anderson, R., & Li, M. (2016). Diagnostic instruction: Towards an integrated system for classroom assessment. In R. A. Duschl & A. S. Bismack (Eds.), *Reconceptualizing STEM education* (pp. 49–67). New York, NY: Taylor Francis/Routledge.
National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. Washington, DC: National Academies Press.
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.
National Research Council. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academic Press.
Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, *50*(2), 162–188.
Nordine, J., Krajcik, J., & Fortus, D. (2011). Transforming energy instruction in middle school to support integrated understanding and future learning. *Science Education*, *95*, 670–699. doi:10.1002/sce.20423
Otero, V. K., & Nathan, M. J. (2008). Preservice elementary teachers' views of their students' prior knowledge of science. *Journal of Research in Science Teaching*, *45*(4), 497–523.
Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching*, *44*(1), 57–84.
Ruiz-Primo, M. A., & Li, M. (2013). Examining formative feedback in the classroom context: New research perspectives. In J. H. McMillan (Ed.), *Handbook of research on classroom assessment* (pp. 215–232). Thousand Oaks, CA: SAGE Publications, Inc.
Shavelson, R., & Kurpius, A. (2012). Reflections on learning progressions. In A. C. Alonzo & A. W. Gotwals (Eds.), *Learning progressions in science: Current challenges and future directions* (pp. 13–26). Rotterdam, The Netherlands: Sense Publishers.
Shavelson, R. J., Young, D. B., Ayala, C. C., Brandon, P. R., Furtak, E. M., Ruiz-Primo, M. A., … Yin, Y. (2008). On the impact of curriculum- embedded formative assessment on learning: A collaboration between curriculum and assessment developers. *Applied Measurement in Education*, *21*(4), 295–314.
Shepard, L. A. (2009). Commentary: Evaluating the validity of formative and interim assessment. *Educational Measurement: Issues and Practice*, *28*(3), 32–37.
Talmy, L. (1985). Force dynamics in language and thought. *Parasession on Causatives and Agentivity*, *1*, 293–337.
Talmy, L. (2000). Toward a Cognitive Semantics, volume I: Concept structuring systems. I–viii, 1–565.
Warren, J. W. (1982). The nature of energy. *European Journal of Science Education*, *4*(3), 295–297.
Watts, D. M. (1983). Some alternative views of energy. *Physics Education*, *18*, 213–217.
Wei, X. (2014). *Developing a learning progression framework for energy in physics across the 8th – 12th grades* (Unpublished doctoral dissertation). Beijing Normal University, China.
Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching*, *46*(6), 716–730.
Yao, J.-X., Guo, Y.-Y., & Neumann, K. (2017). Refining a learning progression of energy. *International Journal of Science Education*, *39*(17), 2361–2381.
Yin, R. K. (2013). *Case study research: Design and methods*. Thousand Oaks, CA: Sage publications.
Yin, Y., Tomita, M. K., & Shavelson, R. J. (2014). Using formal embedded formative assessments aligned with a short-term learning progression to promote conceptual change and achievement in science. *International Journal of Science Education*, *36*(4), 531–552.
Zhai, X., Guo, Y., & Zhang, Y. (2015). *The relationship between teachers' learning progression-based instruction plan determination levels and instruction results – 3 different teachers' case studies*. Proceedings of the Fourth International Conference of East-Asian Association for Science Education (EASE), Beijing. *China*, 38–39.
Zhang, W. (2006). *Textbook for physics of senior high school*. Beijing: People's Education Press.